



Region-based segmentation on depth images from a 3D reference surface for tree species recognition.

Ahlem Othmani, Nicolas Lomenie, Piboule Alexandre, Christophe Stolz, Lew Fock Chong Lew Yan Voon

► To cite this version:

Ahlem Othmani, Nicolas Lomenie, Piboule Alexandre, Christophe Stolz, Lew Fock Chong Lew Yan Voon. Region-based segmentation on depth images from a 3D reference surface for tree species recognition.. IEEE International Conference on Image Processing, Sep 2013, Melbourne, Australia. pp.0003399. hal-00923704

HAL Id: hal-00923704

<https://hal.science/hal-00923704>

Submitted on 3 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

REGION-BASED SEGMENTATION ON DEPTH IMAGES FROM A 3D REFERENCE SURFACE FOR TREE SPECIES RECOGNITION

Ahlem OTHMANI^{1,2}, Nicolas LOMENIE³, Alexandre PIBOULE², Christophe STOLZ¹ and
Lew F.C. LEW YAN VOON¹

¹Université de Bourgogne, Le2i UMR-CNRS 5158, Le Creusot, France

²Office National des Forêts, Département R&D, 54000 Nancy, France

³Université Paris Descartes, LIPADE, Paris, France

ABSTRACT

The aim of the work presented in this paper is to develop a method for the automatic identification of tree species using Terrestrial Light Detection and Ranging (T-LiDAR) data. The approach that we propose analyses depth images built from 3D point clouds corresponding to a 30 cm segment of the tree trunk in order to extract characteristic shape features used for classifying the different tree species using the Random Forest classifier. We will present the method used to transform the 3D point cloud to a depth image and the region based segmentation method used to segment the depth images before shape features are computed on the segmented images. Our approach has been evaluated using two datasets acquired in two different French forests with different terrain characteristics. The results obtained are very encouraging and promising.

Index Terms— Forest inventory, single tree species recognition, depth images from 3D point clouds, depth image segmentation.

1. INTRODUCTION

Airborne and ground remote sensing tools, based on Light Detection and Ranging (LiDAR) technology, are today widely used in the forestry community for forest inventory and management issues. Airborne methods have the advantages of supplying planting level attributes mappings at a forested massif scale. But, they generally lack precision for accurate local estimation, particularly in the case of heterogeneous planting. Consequently, for more accurate estimations of the dendro-metric attributes such as the density of stems, the fine structure of the planting or the site quality, it is necessary to realize forest inventory field plots using T-LiDAR Scanners (TLS) that can produce high resolution three-dimensional (3D) point clouds with centimeter to sub-centimeter precision [1]. Recently, the Office National of Forests ("Office National des Forêts – ONF") has started to develop a software tool named "Computree" for the automatic measurement of forest inventory attributes using TLS data [2]. Among the important attributes that "Computree" should be able to

determine is the species of each tree in the ground plot. The recognition of the species in a mixed planting is essential for many purposes such as, for example, wood volume estimation for each species. But, to the best of our knowledge not much has been done regarding single tree species identification based on TLS data. In this paper we present our work on single tree species identification using only ground plot TLS data. The common criteria used to identify the species of a single tree are the shape of the leaves, the general shape of the crown and the variations in geometry across the surface of the bark known as the geometric texture of the bark. Since forest inventory data are mostly acquired during winter in order to reduce occlusions due to leaves, and mainly so as to do the measurement outside the growing period of the trees, it cannot be envisaged to recognize the species based on the shape of the leaves. Finally, the bark is probably the most discriminating feature of the species even if it is subject to changes during the tree's life because of age, injuries, and modified growth pattern due to environmental disturbances. Figure 1 shows the bark texture of the five most dominant species in European forests. They are the hornbeam, the oak, the spruce, the beech and the pine.



Figure 1: Example of 3D mesh of the five tree species

One can notice that each of the five species has a distinguishable bark feature: the beech has a relatively smooth surface, the spruce is less smooth compared to the beech and it has circular scars, the pine and the oak are rough with vertical strips but the growth pattern is different, and the hornbeam is smooth with an undulating texture. We thus propose a method that analyzes the geometric variation of patterns or geometric texture of the 3D surface of the bark in order to determine the species.

The first step of our method as depicted in section 2 consists in generating a 2D depth map from a 3D point

cloud of a segment of the tree trunk known as a "patch". Then, a region-based depth image segmentation operation is performed so as to extract texture features for the tree species classification step. Section 3 describes the experimental setting and discusses about the results obtained on two datasets, D1 and D2, composed of respectively 16 patches per species and 30 patches per species.

2. METHODOLOGY

Our method consists of four steps as represented in Figure 2. The first step is a denoising step for removing ghost points using the commercially available RapidForm™ software (<http://www.rapidform.com/>). It will not be discussed here.

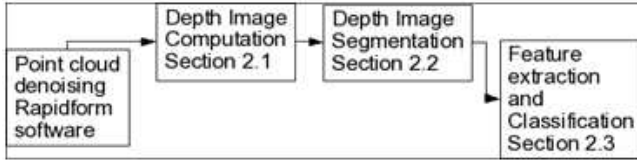


Figure 2: Flow diagram of our method

2.1 Depth image from a reference surface

This step extracts and transforms the geometric details into a depth image. Geometric details are defined as the variations of the geometric distances between the original mesh M_o and a smoothed version M_s as illustrated in Figure 3.

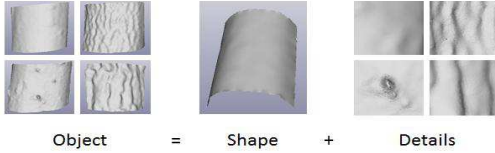


Figure 3 : 3D geometric texture model

Consequently, a depth image of the geometric details can be modeled as

$$f : \mathbb{N}^2 \rightarrow \mathbb{R}$$

$$f(X, Y) = \begin{cases} \text{Mean}(d_i) & \text{if } \exists d_i : \varphi(x_i, y_i, z_i) = (X, Y) \\ \text{bicubic interpolation} & \text{if } \nexists d_i : \varphi(x_i, y_i, z_i) = (X, Y) \end{cases} \quad (1)$$

where

- d_i is the details for a point $v(x_i, y_i, z_i) \in M_s$. d_i is the Euclidean distance between $v(x_i, y_i, z_i)$ and $\tilde{v}(x', y', z')$ the nearest neighbor of $v(x_i, y_i, z_i)$ in the original mesh M_o . $\tilde{v}(x', y', z')$ is determined using the efficient Aligned Axis Bounding Box (AABB) tree structure [3].

$$d_i = d(v, \tilde{v}) = \sqrt{(x_i - x')^2 + (y_i - y')^2 + (z_i - z')^2} \quad (2)$$

$$\tilde{v} = \arg \min_{p_i \in M_s} \|p_i - v\| \quad (3)$$

- $\varphi(\cdot)$ is the Isomap dimensionality reduction algorithm proposed by Tenenbaum et al [4] followed by a quantization step. The latter step is necessary since Isomap transforms nonlinear manifolds in \mathbb{R}^3 to

Euclidean manifolds in \mathbb{R}^2 . A quantization function is thus necessary to map the results of Isomap in \mathbb{R}^2 to the 2D coordinates $(X, Y) \in \mathbb{N}^2$ of the pixels of the depth image.

$$\begin{aligned} \varphi : \mathbb{R}^3 &\rightarrow \mathbb{N}^2 \\ \varphi(x, y, z) &= (X, Y) \end{aligned} \quad (4)$$

In equation (1) the mean of the distances d_i is computed since there may be several 3D points $v(x, y, z)$ that map to the same pixel at coordinates (X, Y) in the depth image. Also, it is possible that no 3D point $v(x, y, z)$ maps to a pixel of coordinates (X, Y) . In that case a bicubic interpolation is done in order to estimate a value for the pixel.

The reference surface must be a sufficiently smoothed mesh where the main structure of the trunk is preserved and all the geometric details have been filtered out. This is achieved using Taubin's λ/μ smoothing algorithm [5] which consists in basically performing the Laplacian smoothing two consecutive times with different scaling factors noted λ and μ . A first step with $\lambda > 0$ (shrinking step) and a second step with a negative scaling factor $\mu < -\lambda < 0$ (unshrinking step). The values of λ and μ are respectively equal to 0.6307 and -0.6732 (values suggested by Taubin). Laplacian smoothing consists in iteratively moving each of the vertices of the mesh to a new position that corresponds to the weighted average position of the neighboring vertices. The new position v'_i of a vertex i is given by:

$$v'_i = v_i + \lambda \Delta v_i \quad (5)$$

where v_i is the current position, λ is a scalar that controls the diffusion speed and Δv_i the Laplacian operator, a weighted sum of the difference between the current vertex v_i and its neighbors v_j , given by the following equation:

$$\Delta v_i = \sum_{j \in i^*} w_{i,j} (v_j - v_i) \quad (6)$$

where i^* is the set of all the neighbors of the vertex v_i . $w_{i,j}$ are the weights that are chosen to be equal for all the neighbors and such that $\sum_j w_{i,j} = 1$.

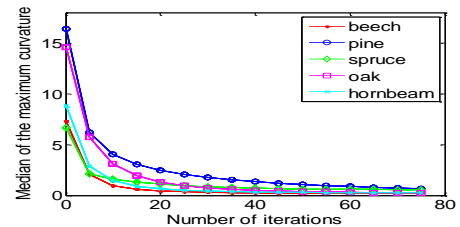


Figure 4: The median of the maximum curvatures against the number of iterations for each of the five tree species

Taubin's algorithm is run iteratively until a smooth surface is obtained. The smoothness of a surface is quantified by the minimum, maximum, mean and Gaussian curvatures of each

of the points of the mesh. A study of the median of these curvatures values as a function of the number of iterations of Taubin's algorithm for several samples of the five species to identify show that there is no significant difference in the curvature against number of iterations curve for the four types of curvature values. We have thus decided to consider only the median of the maximum curvatures curve, as shown in Figure 4, to determine the smoothing stopping criterion. We stop the smoothing process when the slope of the tangent to the curve is less than or equal to -0.01. For this value of the slope we have noticed that the mesh is sufficiently smoothed while preserving the main structures of the trunk. Figure 5 shows some examples of depth images.

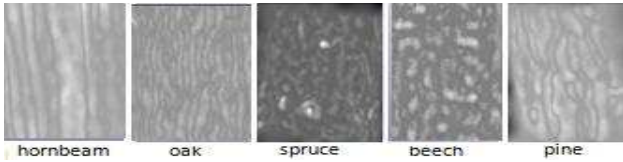


Figure 5: Example of depth image of the five tree species

2.2 Region based segmentation

Tree trunk texture has several discriminating criteria that can be categorized through their degree of roughness which is already present in the depth information: shape (for example, spruce has circular scars) or directionality of the textures (for example, the orientation of the strips is different between hornbeam and pine). Our tree trunk depth images segmentation method should thus highlight the shapes and their directionalities. For that, we propose a region-based segmentation of the depth images which can be described as follows:

Step1: we apply a median filter using a 30×30 window to reduce noise and a contrast stretching to improve the visibility and the contrast of the features. The size of the windows depends on the minimum size of the scars.

Step2: we propose a new filter L5Ed5 adapted to our textures for highlighting the features. We were inspired by the laws filter masks [6]. The measures are derived from three simple vectors. $L3 = [1 \ 2 \ 1]$ that represents averaging and $Ed3 = [-1 \ 0 \ 1]$ that calculates first difference (edges). After convolution of these vectors with themselves, two vectors result: Level $L5 = [1, 4, 6, 4, 1]$ and Edge $Ed5 = [1, 0, -2, 0, 1]$. Multiplying these vectors by considering $L5$ as a column vector and $Ed5$ as a row vector yields a 5×5 Matrix. The resulting filter represents a differentiator kernel of second order on rows, and a smoothing kernel on the columns. It is an image denoising and contrast enhancement filter. We convolve the depth images with the L5Ed5 mask, followed by a moving-window absolute average operation and a normalization by the local mean.

Step3: we repeat step 1 to reduce some blur effect using 15×15 windows for the median filter.

Step 4: segmentation by Otsu thresholding [7].

In Figure 6, an example of a segmented depth image for each of the five species is shown.

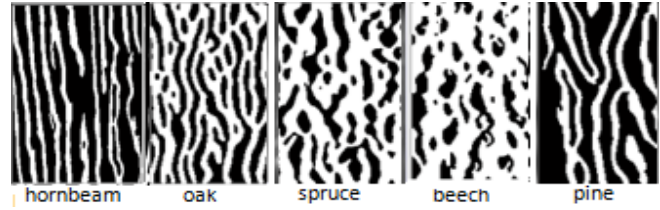


Figure 6: Example of segmented depth images

2.3. Feature extraction and classification

After segmenting the depth images, scars on the trunk will be represented by independent regions. We assign them a set of finite values representing quantitative attributes called features that describes the geometry and the directionality of the regions. We used the `regionprops()` function of Matlab to calculate the properties of the pre-segmented image regions including area, axis length (minor and major), orientation, and so on.

The pre-segmented regions present hollow regions but even salient regions can be very discriminative in some cases (strips size in pines is more important than in the oak). So, we deduce the complement of the segmented image and we compute the shape features of the regions.

Severe injuries may occur during the life of the tree causing considerable damages to the tree bark that change its discriminating criteria. Also, variations of the size of the scars that depend strongly on the age of the tree and the distance to the scanner will lead to a dispersion of the feature values. To evaluate the inter-class variability, we consider the mean, the standard deviation and the median of each feature of all the regions. Classification is done using the Random Forest (RF) classifier proposed by Breiman [8] in order to classify the textures and thus the tree species.

3. EXPERIMENTAL RESULTS AND DISCUSSION

We used two different datasets D1 and D2 to validate our approach experimentally. The test site of D1 is a state mixed forest in Montiers-sur-Saulx, France. The second test site of D2 is a mixed forest stands of grove and coppice-under-grove forests in Lorraine, France.

The T-LiDAR data were collected using a Faro Photon 120™ scanner. They are segmented according to the method described in [2] and 3D patches are extracted to constitute the datasets. Both datasets contain the five different species of trees represented in Figure 1: 16 patches per species in D1 and 30 patches per species in D2. Four tests are done and in each test we train a separate RF classifier built with 1000 decisions trees. For tests 1 and 2, the datasets D1 and D2 are respectively split into 10 equal size subsets and a 10-fold cross-validation process is performed. In tests 3 and 4, D1 is used as the training dataset for testing D2, and vice versa. Confusion tables for the 4 tests are reported in Tables

1, 2, 3 and 4. Table 5 summarizes the min, max and average accuracy of the 4 tests as well as the standard deviation.

	1	2	3	4	5	Acc.
1	11.75	1.916	0.25	0.166	1.916	73.4%
2	0	14.083	1.25	0.25	0.416	88%
3	0	0.083	14.833	1.083	0	92.7%
4	0	0	0.25	15.583	0.166	97.4%
5	0.5	0.833	0.5	0.0833	14.083	88%

Table 1: Confusion matrix for D1 cross validation (1: hornbeam, 2: oak, 3: spruce, 4: beech, 5: pine)

	1	2	3	4	5	Acc.
1	25.666	1.25	0	0.666	2.416	85.6%
2	1.333	28	0	0	0.666	93.3%
3	0	0.083	28.083	1.5	0.333	93.6%
4	1.666	0.416	1.916	25.666	0.333	85.6%
5	1.25	0.416	0	0.583	27.75	92.5%

Table 2: Confusion matrix for D2 cross validation

	1	2	3	4	5	Acc.
1	19	3	0	2	6	63.3%
2	4	26	0	0	0	86.7%
3	0	0	27	3	0	90%
4	3	0	4	23	0	76.7%
5	3	0	0	1	26	86.7%

Table 3: Confusion matrix for training on D1 and testing on D2

	1	2	3	4	5	Acc.
1	10	2	0	0	4	62.5%
2	0	12	3	0	1	75%
3	0	0	14	2	0	87.5%
4	0	0	1	15	0	93.8%
5	0	3	1	0	12	75%

Table 4: Confusion matrix for training on D2 and testing on D1

	Min	Average	Max	σ
Test 1	73.4%	87.9%	97.4%	8.0%
Test 2	85.6%	90.1%	93.6%	3.7%
Test 3	63.3%	80.7%	90%	9.8%
Test 4	62.5%	78.8%	93.8%	10.9%

Table 5: Accuracy rates

We can note that training and testing with dataset D2 (test 2 in Table 5) gives overall better accuracy rate than the other three tests. This may be due to the fact that D2 dataset contains more samples that are better distributed. The standard deviation of the accuracy is also smaller for test 2 meaning that there is less dispersion in the recognition accuracy with respect to the species. Training with one

dataset and testing with the other dataset (tests 3 and 4 in Table 5) gives less good results than training and testing with the same dataset (tests 1 and 2 in Table 5). Nevertheless, the result is relatively good enough and we can consider that the classifier performs reasonably well since the two datasets D1 and D2 are composed of patches from two totally different sites. From the 4 confusion matrices of Tables 1, 2, 3 and 4, one can note that the hornbeam is the species that has the least accuracy rate: an average accuracy of 71.2% over the 4 tests compared to 85% to 91% for the other species.

4. CONCLUSION

In this paper, a new method for tree species recognition with T-LiDAR data has been proposed. The various experiments on two different data sets show that our method has good results (average accuracy over 78.8% with a standard deviation of 10.9% in the worst case and 3.7% in the best case). Future work will focus on finding more efficient features that will better discriminate the species and on the quantification of the performance of our method on other datasets.

5. ACKNOWLEDGEMENT

This work is financially supported by the "Conseil Régional de Bourgogne" under contract N° 2010-9201AAO048S06469 and N° 2010-9201CPERO007S06470, and the "Office National des Forêts".

6. REFERENCES

- [1] M. Dassot, T. Constant and M. Fournier, "The Use of Terrestrial LiDAR Technology in Forest Science: Application fields, Benefits and Challenges", Annals of Forest Science, Vol. 6, pp. 959-974, 2011.
- [2] A. Othmani, A. Piboule, M. Krebs, C. Stolz and L.F.C. Lew Yan Voon, "Towards Automated and Operational Forest Inventories with T-LiDAR", SilviLaser, 2011.
- [3] P. Alliez, S. Tayeb and C. Wormser, "AABB Tree", CGAL 3.5 edition, 2009.
- [4] J.B. Tenenbaum, Vin de Silva and J.C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction", Science, pp. 2319-2323, December 2000.
- [5] G. Taubin, "Geometric Signal Processing on Polygonal Meshes", State of the Art Report, Eurographics, 2000.
- [6] K.I. Laws KI, "Rapid texture identification," Proc. of SPIE Conf. Image Processing for Missile Guidance, Vol. 238, pp. 376-380, 1980.
- [7] Nobuyuki Otsu, "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man and Cybernetics, Vol 9, No.1, pp. 62-66, Jan. 1979.
- [8] L. Breiman, "Random Forests", Machine Learning, pp. 5-32, October 2001.
- [9] Rodrigo de Toledo, Bin Wang, Bruno Levy, "Geometry Textures and Applications," Computer Graphics Forum, Vol. 27, No. 8, pp. 2053-2065, December 2008.